Advances in Mathematics **156**, 157–179 (2000)
doi:10.1006/aima.2000.1940, available online at http://www.idealibrary.com on IDEAL®

Entropy and Dyadic Equivalence of Random Walks on a Random Scenery

Deborah Heicklen

Department of Mathematics, University of California, Berkeley, Berkeley, California 94720-3840; and Lockheed–Martin Corporation

Christopher Hoffman

Department of Mathematics, University of Washington, Seattle, Washington 98195

and

Department of Mathematics, University of Maryland, College Park, Maryland 20742

Received August 14, 1998

For any 1–1 measure-preserving map T of a probability space, consider the $[T, T^{-1}]$ endomorphism and the corresponding decreasing sequence of σ-algebras. We demonstrate that if the decreasing sequence of σ-algebras generated by $[T, T^{-1}]$ and $[S, S^{-1}]$ are isomorphic, then T and S must have equal entropies. As a consequence, if the $[T, T^{-1}]$ endomorphism is isomorphic to the $[S, S^{-1}]$ endomorphism, then the entropy of T is equal to the entropy of S. Central to this is a relationship between Feldman's $\bar{f}$ metric (1976, *Israel J. Math.* **24**, 16–38) and Vershik's v metric (1970, *Dokl. Akad. Nauk SSSR* **193**, 748–751).

1. INTRODUCTION

A decreasing sequence of σ-algebras is a measure space $(X, \mathscr{F}_0, \mu)$ and a sequence of σ-algebras $\mathscr{F}_0 \supset \mathscr{F}_1 \supset \mathscr{F}_2 \cdots$. A natural example of this arises from a sequence of independent identically distributed random variables, $\{X_i\}$. Namely, set $\mathscr{F}_i = \sigma(X_i, X_{i+1}, ...)$. If the X_i take on the values 1 and -1 with probability 1/2, then this sequence has the property that $\mathscr{F}_i \mid \mathscr{F}_{i+1}$ has two point fibers of equal mass for every i. A decreasing sequence of σ-algebras with this property is called *dyadic*. This example also has the

[1] To whom correspondence should be addressed.

0001-8708/00 $35.00

property that $\bigcap \mathcal{F}_i$ is trivial. Two decreasing sequences of σ-algebras are *isomorphic* if there exists a 1–1 measure-preserving map between the two spaces that maps the ith σ-algebras to each other. A dyadic decreasing sequence of σ-algebras is *standard* if it is isomorphic to the decreasing sequence previously mentioned.

Vershik began the modern study of such decreasing sequences of σ-algebras in [10]. In [9] Vershik showed that there exist dyadic sequences of σ-algebras with trivial intersection that are not standard. In [9] Vershik also gave a necessary and sufficient condition for a dyadic decreasing sequence of σ-algebras to be standard. His criterion is more general than for just the dyadic case, but we only mention this case since this is all we use. For a thorough treatment of the classification of decreasing sequences of σ-algebras see [4]. An equivalent description of standardness for dyadic decreasing sequences of σ-algebras is for there to exist a sequence of partitions $\{P_i\}$ of X into two sets, each of measure 1/2, such that

1. the partitions P_i are mutually independent and
2. for each i, $\mathcal{F}_i = \bigvee_{n=i}^{\infty} P_n$.

In this paper we will be working with decreasing sequences of σ-algebras arising from a certain class of endomorphisms. These are known as $[T, T^{-1}]$ *endomorphisms* or random walks on a random scenery. Let T, the scenery process, be any 1–1 measure-preserving map on a probability space $(Y, \mathcal{C}, \nu)$ such that T^2 is ergodic. Let σ be the shift on $(X, \mathcal{B}, \mu)$ where $X = \{-1, 1\}^{\mathbf{N}}$, $\mathcal{B}$ is the Borel σ-algebra, and μ is $(\frac{1}{2}, \frac{1}{2})$ product measure. Define $[T, T^{-1}]$ on $(X \times Y, \mathcal{F}, \mu \times \nu)$ where $\mathcal{F} = \mathcal{B} \times \mathcal{C}$ by

$$[T, T^{-1}](x, y) = (\sigma x, T^{x_0} y).$$

Let $\mathcal{F}_n = [T, T^{-1}]^{-n} \mathcal{F}$. $[T, T^{-1}]$ is 2–1, since any point (x, y) has the preimages $(-1x, Ty)$ and $(1x, T^{-1}y)$. Since each preimage has equal relative measure, the $[T, T^{-1}]$ endomorphism generates a dyadic decreasing sequence of σ-algebras. Furthermore, since T^2 is ergodic, $\bigcap \mathcal{F}_n$ is trivial [8]. Note that if T is the trivial one-point transformation then $[T, T^{-1}]$ reduces to the shift on $(X, \mathcal{B}, \mu)$, and the corresponding sequence is the standard dyadic example.

The above construction, when carried out for $X = \{-1, 1\}^{\mathbf{Z}}$, yields a 1–1 map we refer to as the $[T, T^{-1}]$ *automorphism*. This is also the natural two-sided extension of the $[T, T^{-1}]$ endomorphism. Kalikow proved in [7] that if T has positive entropy then the $[T, T^{-1}]$ automorphism is not isomorphic to a Bernoulli shift. Building on Kalikow's techniques, Heicklen and Hoffman proved that if T has positive entropy then the decreasing sequence of σ-algebras generated by the $[T, T^{-1}]$ endomorphism is not standard [5]. When T has 0 entropy the picture appears to be significantly more complicated. Feldman and Rudolph proved that if T is rank 1 then the $[T, T^{-1}]$ endomorphism generates a standard decreasing sequence of σ-algebras [3]. On the other hand, Hoffman has given an example of a zero entropy T such that the $[T, T^{-1}]$ endomorphism is not standard [6]. For most zero-entropy transformations it is not known whether the $[T, T^{-1}]$ endomorphism generates a standard decreasing sequence of σ-algebras. Feldman and Rudolph's result, combined with the work of Burton [1], provides an example of an endomorphism which is standard but is not isomorphic to a Bernoulli shift.

The purpose of this paper is to demonstrate that if the decreasing sequence of σ-algebras generated by $[T, T^{-1}]$ and $[S, S^{-1}]$ are isomorphic, then the entropy of T is equal to the entropy of S. Since two endomorphisms which are isomorphic produce isomorphic decreasing sequences of σ-algebras, this result implies that if the $[T, T^{-1}]$ endomorphism is isomorphic to the $[S, S^{-1}]$ endomorphism then the entropy of T is equal to the entropy of S. It is not known if there exist two transformations S and T with different entropies such that the $[T, T^{-1}]$ automorphism is isomorphic to the $[S, S^{-1}]$ automorphism. Our methods rely heavily on the tree structure of an endomorphism, which is explained shortly. As an isomorphism of the $[T, T^{-1}]$ automorphism and the $[S, S^{-1}]$ automorphism need not preserve the tree structure of the endomorphism our methods do not apply to the invertible case.

There is a tree structure associated with the preimages of any point in the $[T, T^{-1}]$ endomorphism. We introduce some notation for this tree structure. Fix a value n for the height of the tree. An *n branch* is an element $b \in \{-1, 1\}^n$. An *n tree* is a binary tree of height n consisting of 2^n branches. Given a point (x, y), there is a map associating the 2^n branches of the n tree with the 2^n elements of $[T, T^{-1}]^{-n}(x, y)$. This map takes the n branch $b = (b_1, ..., b_n)$ to $(-b_1, ..., -b_n x, T^{\sum b_i} y)$. Based on this, if P is a finite partition of Y the *labeled n tree* for a partition P over a point $y \in Y$ assigns to each branch b the label $P(T^{\sum b_i} y)$.

For $m \leqslant n$ define an *m subtree inside an n tree* to be a tree with 2^m branches such that the last $n-m$ coordinates of the branches all agree and the first m coordinates vary over all possibilities. The 2^m branches of an m tree inside an n tree are associated with the 2^m elements of $[T, T^{-1}]^{-n}(x, y)$ which are mapped to the same point under $[T, T^{-1}]^m$.

For our purposes a *node* is an integer. We will say that a branch *lands at a node k* if $\sum_1^n b_i = k$. We will say a branch *passes through a node k at height $h > 1$* if $\sum_h^n b_i = k$. This vocabulary will be used repeatedly in Section 3.

The *distance between any two m subtrees in an n tree* is $|\sum_{i=m+1}^n b_i - b'_i|$ for any branches b and b' passing through the two subtrees. Fix a partition

P and let W and W' be the labeled n trees over y and y' respectively. The Hamming metric between two labeled n trees W and W' is given by

$$d_n(W, W') = \frac{\#\text{ of branches on which the labels of } W \text{ and } W' \text{ disagree}}{2^n}.$$

Let $\mathscr{A}_n$ be the group of automorphisms of an n tree. For an automorphism $a \in \mathscr{A}_n$, let $f(a)$ be the largest integer such that $(a(b))_i = b_i$ for all $1 \leqslant i \leqslant f(a)$ and all branches b if such an integer exists. If no such integer exists set $f(a) = 1$. Define $g(a) = 1/(1 + \log f(a))$ if $a \neq id$ and $g(a) = 0$ if $a = id$.

Define

$$v_n^P(y, y') = \inf_{a \in \mathscr{A}_n} (d_n(aW, W') + g(a)).$$

This is easily checked to be a pseudometric on points y and y' as it is a metric on the labeled n trees W and W'. If the partition P is understood, it will be omitted from the formulation and we will write v_n.

In Section 2 we will briefly outline the path to the proof of our main result. In Sections 3 and 4 we prove that a tree automorphism on a labeled tree which produces a small minimum for distance $v_n(y, y')$ must have a certain form. This will establish a connection between v_n and Feldman's $\bar{f}$ metric. In Section 5 we use this property in finite code approximations to Φ to prove that if the $[T, T^{-1}]$ and $[S, S^{-1}]$ endomorphisms generate isomorphic sequences of σ-algebras then T and S have the same entropy by showing that the exponential growth rate for T-names will bound that for S-names.

2. OUTLINE OF THE PROOF

On the way to proving our main result we answer the following two questions:

1. What kind of automorphism $a \in \mathscr{A}_n$ will minimize the value of v_n between the labeled trees over two points?
2. Given T and a sufficiently small δ, how does $v\{y' \mid v_n^P(y, y') < \delta\}$ behave as $n \to \infty$?

The answers to these questions will establish a connection between v_n and Feldman's $\bar{f}$ metric between the T, P, n-names of y and y'. We begin with the following result from [5].

THEOREM 2.1 [5]. *If T is a positive entropy transformation then there exists δ_0 and a finite partition P such that (T, P) is an i.i.d. process and for any polynomial $p(n)$, n sufficiently large, and any point y;*

$$\nu\{y' \mid v_n^P(y, y') < \delta_0\} < \frac{1}{p(n)}.$$

In the case of $[T, T^{-1}]$ endomorphisms Vershik's standardness criteria says that

$$\int v_n^P(y, y')\, d\nu \times \nu \to 0$$

for every finite P iff $\{\mathscr{F}_n\}$ is standard [9]. Thus Theorem 2.1 implies that if T has positive entropy then the decreasing sequence of σ-algebras generated by the $[T, T^{-1}]$ endomorphism is not standard.

We strengthen this result to show that for δ small enough, $\nu\{y' \mid v_n^P(y, y') < \delta\}$ decays exponentially in $\sqrt{n}$. The proof of Theorem 2.1 in [5] does not give any indication of what type of tree automorphism is used to obtain $v_n^P(y, y') < \delta_0$. Building from the conclusion of this theorem we show that such an automorphism must take on a certain form. Namely, for most nodes k, it maps almost all of the branches that land at k to branches that land at some single node k'. This gives a 1–1 map from most values k to the value $A(k) = k'$. In showing this we will further establish that on a large subset of nodes the map A is monotone.

This map A will establish a connection between v_n and the $\bar{f}$ of Feldman [2]. To remind the reader, the $\bar{f}$ metric is defined as follows. For any m, n, and $w, w' \in \{0, ..., l\}^{\mathbf{Z}}$ let

$$\bar{f}_{[m,n]}(w, w') = 1 - \frac{k}{n-m+1},$$

where k is the maximal integer for which there are subsequences of integers, $m \leqslant i_1 < i_2 < \cdots < i_k \leqslant n$ and $m \leqslant j_1 < j_2 < \cdots < j_k \leqslant n$ such that $w(i_r) = w'(j_r)$, $1 \leqslant r \leqslant k$. We also will use the $\bar{d}$ metric on sequences. For any m, n, and $w, w' \in \{0, ..., l\}^{\mathbf{Z}}$ let

$$\bar{d}_{[m,n]}(w, w') = 1 - \frac{k}{n-m+1},$$

where k is the number of i, $m \leqslant i \leqslant n$, such that $w(i) = w'(j)$.

There are some difficulties to expressing this connection simply. For one, if n is even then $v_n(y, y')$ depends only on the even coordinates of y and y' between $-n$ and n, while $\bar{f}$ depends on all the coordinates. Also, the number of branches in the n tree over y landing at node k is given by the binomial distribution. Thus $v_n(y, y')$ depends much more heavily on the values of y_i and y'_i for $|i| < \sqrt{n}$ than for $|i| > \sqrt{n}$. On the other hand, $\bar{f}_{[-n, n]}$ gives uniform weight to all coordinates in the interval $[-n, ..., n]$. Hence if $v_n(y, y')$ is small, then we can only draw conclusions about $\bar{f}_{[-c\sqrt{n}, c\sqrt{n}]}(y, y')$. To overcome the first of these difficulties (and others as well) we assume that the generating partition P has a certain form. Let Q be a partition which generates and is a refinement of a full entropy and i.i.d. partition for the action of T. Now set $P = Q \vee T(Q)$. Hence the sequence $T^i(P)$ for i even determines $T^i(Q)$ for all i. We also will restrict ourselves to values n that are even and perfect squares. We obtain the following relationship between v_n and $\bar{f}_{[-c\sqrt{n}, c\sqrt{n}]}$.

THEOREM 2.2. *Assume n is even and a perfect square and that P is as described above. There is a constant C so that given ε and $c > 1$ there exist $\delta > 0$ and a good set G with $\mu(G) > 1 - \varepsilon$, and for $y, y' \in G$, if $v_n(y, y') < \delta$ then once n is large enough*

$$\bar{f}_{[-c\sqrt{n}, c\sqrt{n}]}(y, y') < C/\sqrt{c}.$$

The proof of this result is developed in Section 3 and completed in Section 4; except for certain technical material that appears in the last of Section 6. In Section 5 we use finite code approximations to the isomorphism of the decreasing sequence of σ-algebras and Theorem 2.2 to show that entropy is an invariant. This will require explicit understanding of how δ depends on c in Theorem 2.2. Ignoring sets of small measure this is done as follows. Let Φ be the isomorphism between $[T, T^{-1}]$ and $[S, S^{-1}]$ and suppose that $\Phi(x, y) = (w, z)$ and $\Phi(x', y) = (w', z')$. We will show that this implies that z and z' are close in $\bar{f}$. This implies that the number of names in the T process must grow at the same exponential rate as the number of names in the S process. Hence the entropies of the two processes must be equal.

3. TREE AUTOMORPHISMS

In this section we will show that for most y and y', if $v_n(y, y')$ is small enough then an automorphism a which achieves this small value induces a bijection A from most of $[-n, ..., n]$ to most of $[-n, ..., n]$ with the following property. For most of the branches which land at node k in the tree over y, the image of those branches under A lands at the node $A(k)$ in the tree over y'.

First we define the small set of points y that are *degenerate*. We will deal only with points that are not degenerate. Then we define what it means for a branch in a tree to be *very good*. The very good branches are the large set of branches mentioned in the previous paragraph. Finally, in Lemmas 3.5 and 3.6, we will show how the automorphism a induces the function A.

For the rest of this section and the next we will set and then work with two parameters, δ_0 and δ. At this point we assume δ_0 to be a value satisfying Theorem 2.1. After Lemma 3.1 we will fix δ_0. After Lemma 3.2 we will fix δ. Given a value for δ, fix $n_0 = \lfloor 2^{1/\delta} \rfloor$, where $\lfloor x \rfloor$ means the greatest integer less than or equal to x. Remember that we have fixed a partition P of the space Y. The metric v_k depends on P but we suppress this dependence in the notation. Now we restrict our attention to a certain (large) class of y and y'. For $y \in \{0, ..., l\}^{\mathbf{Z}}$, define

$$\gamma_k = \{ y \mid \exists i \; 2k < |i| < k^5 \text{ such that } v_k(y, T^i(y)) < \delta_0 \}.$$

Also define

$$B(\delta, \delta_0) = \{ y \mid \exists i \; 0 < |i| < n_0 \text{ such that } \bar{d}_{[-n_0, n_0]}(y, T^i(y)) < \delta_0/4 \}.$$

Given a point y we say a node l is *colored* if

1. there exists j and k such that $T^j(y) \in \gamma_k$ and $l \in [j-k, j+k]$ or
2. there exists j such that $T^j y \in B(\delta, \delta_0)$ and $l \in [j-n_0, j+n_0]$.

We will use the term, "the n *word*" of y to refer to the string of symbols $y_{-n}, ..., y_n$, where $y_i = P(T^i(y))$. The colored nodes are bad for our arguments and our immediate goal is to show that they are scarce. The translation interval $[2k, k^5]$ in the definition of γ_k is set as such to ensure that the measure of γ_k is small. A lower bound on the amount of translation is needed because any word can be matched pretty well in the v_n metric to a small translate of itself. See [3] for how this can be achieved. An upper bound is needed since if we allow a significantly larger translate (i.e., exponential in k), then any word is bound to eventually reoccur. Thus any word will have a close $\bar{d}$ matching (and a close v_n matching) to some translate of itself if the window size is large enough.

DEFINITION 3.1. The n word of y is (δ, δ_0) *degenerate* if the fraction of branches in the n tree over y that land at colored nodes is more than $\delta/2$.

The simplest example of a degenerate word is one with all symbols equal. We want to rule out this type of word, since applying any automorphism to this word will not change it and such an arbitrary automorphism does not have the property described earlier of acting essentially as a permutation on the nodes.

LEMMA 3.1. *Depending only on the process* (T, P), *there exists* δ_0 *such that for all* δ *sufficiently small*, $\mu(B(\delta, \delta_0)) < \delta^2/4$.

Proof. Suppose $\{y_i\}$ are i.i.d. random variables on symbols $\{1, ..., \ell\}$ with probability distribution $\{p(1), ..., p(\ell)\}$. Assume $p(1)$ to be the largest term of the probability vector. For any fixed i, in order to get a $\bar{d}$ matching to within $\delta_0/4$, at least a fraction $1 - \delta_0/4$ of the coordinates must match. There are

$$\binom{2n_0+1}{(2n_0+1)\frac{\delta_0}{4}}$$

different choices of subsets of indices to match. The probability that any particular pair matches is less than $p(1)^{(1-\delta_0/4)(2n_0+1)}$. Hence

$$\mu(\{y \mid \bar{d}_{[-n_0, n_0]}(y, T^i(y)) < \delta_0/4\}) \leqslant \binom{2n_0+1}{(2n_0+1)\frac{\delta_0}{4}} p(1)^{(1-\delta_0/4)(2n_0+1)}.$$

Using an estimate on the binomial coefficients and summing over all possible i yields

$$\mu(B(\delta, \delta_0)) \leqslant 2n_0 2^{h(\delta_0/4)(2n_0+1)} p(1)^{(1-\delta_0/4)(2n_0+1)}.$$

In our case P refines a fixed i.i.d. process and so we obtain this bound on $\mu(B(\delta, \delta_0))$ where $p(1)$ is the probability of the largest symbol in the i.i.d. process. We assume that δ_0 has been chosen small enough that

$$p(1)^{(1-\delta_0/4)} < 2^{-2h(\delta_0/4)}$$

and hence

$$\mu(B(\delta, \delta_0)) \leqslant 2n_0 2^{-h(\delta_0/4)(2n_0+1)}.$$

Since $n_0 > 2^{1/\delta} - 1$, $\mu(B(\delta, \delta_0))$ decays superexponentially in δ. Hence, for δ sufficiently small $\mu(B(\delta, \delta_0)) < \delta^2/4$. ∎

For the rest of our work we fix δ_0 such that the previous lemma is satisfied.

LEMMA 3.2. *For any sufficiently small δ, the set of all y whose n-word is degenerate for a given $n > n_0 = \lfloor 2^{1/\delta} \rfloor$ has measure less than δ.*

Proof. Applying Theorem 2.1 with $p(k) = k^{11}$ shows that if δ is small enough then there is an n_0 with $v(\gamma_k) < 1/k^6$ for $k > n_0$. Thus if δ is small enough then $\sum_{k \geqslant n_0} 2kv(\gamma_k) < \delta^2/4$. Using Lemma 3.1 we can choose δ small enough so that $\mu(B(\delta, \delta_0)) < \delta^2/4$. Now Chebychev's lemma implies that the set of all y which are degenerate has measure less than δ. ∎

Now we fix δ so that the previous two lemmas are satisfied and so that $n_0 = \lfloor 2^{1/\delta} \rfloor > N_0$, where N_0 is defined in Lemma 6.4. Theorem 2.2 and the first paragraph in this section apply to all points with nondegenerate n words. In particular, we can choose the good set G in this theorem to be the set of points whose n words are nondegenerate. For the rest of this section and the next section we will fix, in addition to δ and δ_0, n, two points y and y' whose n words are not degenerate, and $v_n(y, y') < \delta$, and an automorphism $a \in \mathscr{A}_n$ that minimizes the quantity in the definition of $v_n(y, y')$.

Now we want to define the large set of very good branches on which the automorphism acts well. Before we can do this we need one intermediate definition.

DEFINITION 3.2. A branch b is *good* for the automorphism a if the following hold:

1. The label assigned to b in the n tree over y is the same as the label assigned to $a(b)$ in the n tree over y',
2. b and $a(b)$ do not land on a colored node, and
3. b and $a(b)$ are n_0 regular.

We defer defining the large set of "n_0 regular" branches until Section 6. The property of two n_0 regular branches that we will use is if b and b' are n_0 regular branches and $g \geqslant n_0$, then there exists an $h \in (g^{1/4}, g^{1/2})$ such that $|\sum_1^h b_i|, |\sum_1^h b_i'| < 101\sqrt{h}$. The difficulty arises in finding a height h that satisfies this for both b and b' simultaneously. This property of n_0 regular branches is proved in Lemma 6.5.

LEMMA 3.3. *If the n words of y and y' are not degenerate and $v_n(y, y') < \delta$, then the fraction of good branches is greater than $1 - 10\delta$.*

Proof. Because $v_n(y, y') < \delta$ the fraction of branches satisfying the first property is greater than $1-\delta$. Since y and y' are not degenerate the fraction of branches satisfying the second property is greater than $1-\delta$. In Lemma 6.4 we will show that the fraction of branches satisfying the third property is more than $1-8\delta$. ∎

DEFINITION 3.3. A branch b is *very good* for the automorphism a if for every $j \leqslant n$ the j subtree through which the branch b passes has at least $2^j(1-C_1)$ good branches, where $C_1 = ((\delta_0/16)(\int_{202}^{\infty} e^{-x^2/2} dx))$.

LEMMA 3.4. *If the n words of y and y′ are not degenerate, and* $v_n(y, y') < \delta$, *then the fraction of very good branches is more than* $1-(10/C_1)\,\delta$.

Proof. For each branch b which is not good, consider the maximal j such that the j subtree containing b has less than $2^j(1-C_1)$ good branches. For this subtree, the fraction of branches in this j subtree which are not good is greater than C_1. Hence

$$\frac{\#\text{ of branches in this subtree}}{\#\text{ of not good branches in this subtree}} \leqslant \frac{1}{C_1}.$$

The union of all these subtrees is the set of all not very good branches. This is because we chose j to be maximal. Hence any branch which is not contained in one of the previously mentioned subtrees always passes through nodes with a more than $1-C_1$ fraction of good branches. Thus the fraction of branches which are not very good is less than $10\delta/C_1$. ∎

We next show that the automorphism a maps all of the very good branches in the tree over y that pass through one node k at height n_0 to branches in the tree over y' that also pass through a single node. In fact, the image of two very good branches that pass through the same node at height $h \geqslant n_0$ will be separated by no more than a distance $2h$. Before stating and proving the result formally, we give an idea of the proof.

Since the two branches pass through the same node at height $h \geqslant n_0$, the images of the two h subtrees containing the two very good branches can be matched well in the Hamming distance after some tree automorphism. Thus the y' names in the intervals that the trees lie over are close in v_n. This implies that either the intervals overlap or they are separated by at least h^5. If they are far apart then, at some height h', the images of the two very good branches must have been separated by some distance between $2h'$ and $(h')^5$. We can use the regularity condition to show that this implies there must be some good branches landing on colored nodes, which is a contradiction.

To bootstrap our way we will need to consider here not only very good branches that pass through a common node, but also ones that pass within 4 of a common node at some height $h \geqslant n_0$.

LEMMA 3.5. *Suppose the n words of y and y′ are not degenerate and $v_n(y, y') < \delta$. Given b and b′, any two very good branches for a, and $h \geqslant n_0$, if*

$$\sum_{i=h+1}^{n} b_i - b'_i \leqslant 4 \qquad \textit{then} \quad \left| \sum_{i=h+1}^{n} a(b)_i - a(b')_i \right| \leqslant 2h.$$

Proof. The proof is by contradiction. Suppose there exist two very good branches, b and b', and height $h \geqslant n_0$ such that b and b' pass through the nodes at most 4 apart at height h, but $a(b)$ and $a(b')$ are in h subtrees that are at least $2h$ apart. Call these subtrees $t_{a(b),h}$ and $t_{a(b'),h}$. Since b and b' are very good they do not land on colored nodes, and $t_{a(b),h}$ and $t_{a(b'),h}$ must be a distance $g \geqslant h^5$ apart as they are within δ_0 in v_h. By Lemma 6.5 it is possible to find a height h' with $g^{1/4} < h' < g^{1/2}$ such that b and b' are in h' subtrees, $t_{b,h'}$ and $t_{b',h'}$, that are less than $202\sqrt{h'}$ apart.

Define the *overlap* between $t_{b,h'}$ and $t_{b',h'}$ at a node j at height $h'/4$ to be the minimum of the fraction of branches in $t_{b,h'}$ that pass through node j at height $h'/4$ and the fraction of branches in $t_{b',h'}$ that pass through node j at height $h'/4$. The overlap summed over all nodes at height $h'/4$ between the two subtrees is greater than $\frac{1}{4}(\int_{202}^{\infty} e^{-x^2/2}\,dx)$. Since the fraction of branches in $t_{b,h'}$ and $t_{b',h'}$ that are not good is small compared with this number, there exists a choice of node J and two $h'/4$ subtrees, $t_{\tilde{b},h'/4}$ and $t'_{\tilde{b}',h'/4}$, in $t_{b,h'}$ and $t_{b',h'}$ respectively, which pass through the node J at height $h'/4$ and have at least a $1-\delta_0/2$ fraction of good branches. Moreover, the distance between the images of these two subtrees, $t_{a(\tilde{b}),h'/4}$ and $t_{a(\tilde{b}'),h'/4}$, is between $h'/2$ and $(h'/4)^5$. Now the labeled trees corresponding to $t_{\tilde{b},h'/4}$ and $t_{\tilde{b}',h'/4}$ are the same. The fraction of good branches in $t_{\tilde{b},h'/4}$ and $t_{\tilde{b}',h'/4}$ is large enough so that

$$v_{h'/4}(t_{\tilde{b},h'/4}, t_{a(\tilde{b}),h'/4}) \leqslant \delta_0/2 \qquad \text{and} \qquad v_{h'/4}(t_{\tilde{b}',h'/4}, t_{a(\tilde{b}'),h'/4}) \leqslant \delta_0/2.$$

Thus

$$v_{h'/4}(t_{a(\tilde{b}),h'/4}, t_{a(\tilde{b}'),h'/4})) \leqslant \delta_0.$$

This implies that all of the branches in $t_{a(\tilde{b}),h'/4}$ and $t_{a(\tilde{b}'),h'/4}$ land on colored nodes. This is a contradiction with the fact that most of these branches are good. ∎

LEMMA 3.6. *If the n words around y and y′ are not degenerate and* $v_n(y, y') < \delta$, *then for any two very good branches b and b′ passing through nodes k and k′ at height* n_0 *with* $|k-k'| \leqslant 4$ *we have* $a(b)$ *and* $a(b')$ *passing through nodes* $\tilde{k}$ *and* $\tilde{k}'$ *at height* n_0 *with* $k-k' = \tilde{k} - \tilde{k}'$. *In particular, all very good branches that pass through a single node at height* n_0 *are mapped to a set of branches also all passing through a single node at height* n_0.

Proof. We need to show that if b and b' are two very good branches and $\sum_{n_0+1}^h b_i = k$ and $\sum_{n_0+1}^h b'_i = k'$ with $|k-k'| \leqslant 4$ then $\sum_{n_0+1}^h a(b)_i = \sum_{n_0+1}^h a(b')_i + k' - k$. By the previous lemma

$$\left| \sum_{n_0+1}^{h} a(b)_i - \sum_{n_0+1}^{h} a(b')_i \right| \leqslant 2n_0.$$

As $f(a) \geqslant n_0$ we have that $y'_{\sum_{n_0+1}^h a(b)_i - n_0}, \ldots, y'_{\sum_{n_0+1}^h a(b)_i + n_0}$ and $y'_{\sum_{n_0+1}^h a(b')_i + k' - k - n_0}, \ldots, y'_{\sum_{n_0+1}^h a(b')_i + k' - k + n_0}$ are within $\delta_0/4$ in $\bar{d}$. If

$$\sum_{n_0+1}^{h} a(b)_i - \sum_{n_0+1}^{h} a(b')_i \neq k - k'$$

then $a(b)$ and $a(b')$ land on colored nodes. This is a contradiction. ∎

Define $A(k) = k'$ for any node k at height n_0 through which a very good branch passes. It remains to be seen why the behavior of A at height n_0 descends to height 0. This will be done in the next section, after verifying that A is monotone on most nodes.

4. THE RELATION BETWEEN v_n AND $\bar{f}_{[-c\sqrt{n},\, c\sqrt{n}]}$

In this section we complete the proof of Theorem 2.2 based on the conclusions of the previous section. We still have fixed the same y, y', a, δ, and A from the previous section. We also fix $c > 0$. We want to show that A is monotone on a large part of the interval $[-c\sqrt{n}, c\sqrt{n}]$. This will tell us that y and y' are $\bar{f}$ close on this interval.

First we define the set of *good* nodes. We prove that the density of good nodes in the interval $[-c\sqrt{n}, c\sqrt{n}]$ is close to 1. In Lemma 4.2 we show that if j and k are good nodes and $|j-k|$ is small then the ratio of $|j-k|$ to $|A(j) - A(k)|$ is close to one. These two facts allow us to define a set of nodes of slightly smaller density on which A is monotone. Once we know that A is monotone on a set of large density by requiring that $v_n(y, y') < \delta/10$, that is to say somewhat smaller than δ, it will follow that A descends to most nodes at height 1 and that $\bar{f}(y, y')$ is small.

DEFINITION 4.1. A node k at a height h is *good* for the automorphism a if a fraction at least $1-C_2\sqrt{c}$ of the branches with $\sum_{i=h+1}^{n} b_i=k$ are very good, where $C_2=\frac{1}{4}(\int_{1/2}^{1} e^{-x^2/2}\,dx)(\frac{1}{2}e^{-1/8})$.

LEMMA 4.1. *For $F(c)=40\,e^{c^2/2}/C_1C_2\sqrt{c}$ at least a $1-F(c)\,\delta$ fraction of the nodes in $[-c\sqrt{n}, c\sqrt{n}]$ at height n_0 are good.*

Proof. Each of these nodes has at least $\frac{1}{4}(2^{n-n_0}/\sqrt{n-n_0})\,e^{-c^2/2}$ branches landing at this node. Thus for each bad node in this region there are at least $(C_2)\,\frac{1}{4}(2^{n-n_0}/\sqrt{n-n_0})\,e^{-c^2/2}$ branches that are not very good. Thus, by Lemma 3.7, there are at most

$$\frac{2^{n-n_0}(10\delta)/C_1}{(C_2)(2^{n-n_0}e^{-c^2/2})/4\sqrt{n-n_0}}<\frac{40\,e^{c^2/2}}{C_1C_2}\,\delta\sqrt{n}$$

nodes which are not good. ∎

We will now use this lemma to prove that A does not distort the distances between two good nodes at height n_0 which are close together by too much.

LEMMA 4.2. *Given c and nodes k and j at height n_0 that satisfy*

1. *k and j are good,*
2. *$|k-j|<.1\sqrt{n}$, and*
3. *$|k|, |j|<c\sqrt{n}$,*

then $.5\,|k-j|\leqslant|A(k)-A(j)|\leqslant 2\,|k-j|$.

Proof. The local central limit theorem tells us that for a fixed c there exists an N such that for any $n\geqslant N$, $|k-j|<.1\sqrt{n}$, and $|k|, |j|<c\sqrt{n}$, the minimum of the two distributions

$$P\left(\sum_{i=n_0+1}^{(j-k)^2} b_i \,\middle|\, \sum_{i=n_0+1}^{n} b_i=k\right) \quad\text{and}\quad P\left(\sum_{i=n_0+1}^{(j-k)^2} b_i \,\middle|\, \sum_{i=n_0+1}^{n} b_i=j\right)$$

has an amount of mass that is bounded below by $\frac{1}{4}(\int_{1/2}^{1} e^{-x^2/2}\,dx)$.

Now, in each one of these subtrees in the overlap, if we have a subtree of height $j-k$ centered over c at height n_0, $k\leqslant c\leqslant j$, and $|A(k)-A(j)|\geqslant 2\,|k-j|$, then we have two possibilities. Either $A(j)$ is at least $(j-k)/2$ further away from the center of the image of the subtree over c than j was from the center of the subtree over c, or the same statement applies with $A(k)$ and k replacing $A(j)$ and j. Without loss of generality assume that this applies to j. Then the ratio of the number of branches in the image of the subtree that pass through $A(j)$ to the number of branches in the subtree that pass through j is at most $1-\frac{1}{2}e^{-1/8}$. Thus the fraction of branches in the subtree that passes through j but whose image is not at $A(j)$ is greater than $\frac{1}{2}e^{-1/8}$. Now a fraction of at least $\frac{1}{4}(\int_{1/2}^{1} e^{-x^2/2}\,dx)(\frac{1}{2}e^{-1/8})$ of the branches that pass through j do not pass through $A(j)$. Since this applies to all subtrees centered between j and k, this contradicts the way C_2 was chosen. ∎

Now we will restrict to a smaller set of nodes on which A is monotone.

Definition 4.2. Let K be the set of all nodes k at height n_0 such that

1. $k \in [-c\sqrt{n}, c\sqrt{n}]$,
2. k is even,
3. there does not exist an interval $[i, j]$, $-c\sqrt{n} \leqslant i \leqslant k \leqslant j \leqslant c\sqrt{n}$, where at least $\frac{1}{4}$ of the even nodes in $[i, j]$ are not good.

Lemma 4.3. *If* $v_n(y, y') \leqslant (1/160c^{1.25}F(c))$ *then* $|K^c| < \sqrt{n}/20c^{.25}$ *and* A *is monotone on* K.

Proof. By Lemma 4.1 a fraction of at least $1-F(c)\,\delta$ of the nodes in the interval are good. Restricted to this interval, K^c can be written as the union of all intervals $[i, j] \subset [-c\sqrt{n}, c\sqrt{n}]$ such that more than $\frac{1}{4}$ of (i, j) is bad. There exists a disjoint collection of such intervals which covers at least half of K^c. Thus $|K^c| \leqslant 8F(c)\,c(1/160c^{1.25}F(c))\sqrt{n} \leqslant \sqrt{n}/20c^{.25}$.

It suffices to show that A is monotone on the intersection of K and any interval of length $.1\sqrt{n}$ between $-c\sqrt{n}$ and $c\sqrt{n}$. This is true because the previous paragraph assures us that in each one of these intervals there are at least two nodes in K. Choose any of these intervals. Take $k \in K$ in this interval. Let k' be the minimal good node inside this interval such that $k' > k$ and $A(k') < A(k)$. By Lemma 3.6 $|k-k'| > 4$. Find a good node j such that $j \in (k, k')$ and $(k'-k) > 4(k'-j)$. Then

$$\begin{aligned} A(j) - A(k') &= A(j) - A(k) + A(k) - A(k') \\ &> A(k) - A(k') \geqslant .5(k'-k) > 2(k'-j). \end{aligned}$$

This contradicts Lemma 4.2. ∎

Assume now that $v_n(y, y') < \delta/10$ where $\delta < 1/160c^{1.25}\,F(c)$ and the tree automorphism a achieves this v_n distance. Recall that the parameters δ and n_0 are linked by the relation $n_0 = \lfloor 2^{1/\delta} \rfloor$. Hence we can apply all our reasoning simultaneously now for the two values δ and $\delta_1 = \log(400n_0^2)^{-1} > \delta/10$ and hence at heights n_0 and $400n_0^2$. To each of these choices will correspond sets of very good branches. We will refer to these as very good branches for δ and δ_1 respectively.

DEFINITION 4.3. For $v_n(y, y') < \delta/10$ let K_1 be the set of all nodes $k \in K$ (the set K for δ) such that k is also good relative to δ_1.

One concludes from Lemmas 4.3 and 4.1 that $|K_1^c| \leqslant \sqrt{n}/10c^{.25}$.

LEMMA 4.4. *If y and y' are not degenerate and $v_n(y, y') < \delta/10$ where $\delta \leqslant 1/160c^{1.25} F(c)$ then for any $k \in K_1$, A acts as a translation on the set of nodes $K_1 \cap [k-2n_0, k+2n_0]$.*

Proof. If $k' \in k_1 \cap [k-2n_0, k+2n_0]$ then all but a fraction $2C_2\sqrt{c} < 2C_2$ of the branches through k and through k' are very good for both δ and δ_1. There must then be two such branches b and b' passing through k and k' respectively at height n_0 and passing through the same node t at height $400n_0^2$. It now follows from Lemma 3.6 applied to δ_1 that $a(b)$ and $a(b')$ must pass through the same node at height $400n_0^2$. They also must pass through $A(k)$ and $A(k')$ at height n_0. As $f(a) > 100n_0^2$, i.e., the tree automorphism a preserves the shape of branches up to height $100n_0^2$, we have the result. ▮

At this point our work is asymmetric but as $v_n(y, y') < \delta/10$ is equivalent to $v_n(y', y) < \delta/10$ we can draw identical conclusions about a^{-1} as we do for a. That is to say there is a set of very good branches in the tree labeled by y' relative to a^{-1} and a set of nodes K_1' at height n_0 in the tree labeled by y' exhibiting the symmetric properties of K_1 but for the tree automorphism a^{-1}.

DEFINITION 4.4. We say that a branch b is *nice* if it is very good relative to a in the tree over y and passes through a node of K_1 at height n_0 and its image branch $a(b)$ has the symmetric properties of being very good relative to a^{-1} and passing through a node of K_1' at height n_0. Let K_2 be the set of all nodes at the base of the tree that have at least one nice branch landing on them.

COROLLARY 4.1. *Suppose y and y' are nondegenerate and $v_n(y, y') < \delta/10$. Further suppose that b and b' are nice branches landing at k and k' respectively and $a(b)$ and $a(b')$ land at $\tilde{k}$ and $\tilde{k}'$ respectively. Then $k \geqslant k'$ iff $\tilde{k} \geqslant \tilde{k}'$. Hence $\tilde{k}$ depends only on $k \in K_2$ and not on the choice of the nice branch b and there is a permutation $\tilde{A}$ of K_2 with $\tilde{k} = \tilde{A}(k)$.*

Proof. As $f(a) > n_0$, if this were not true then for one of the pairs of branches b, b' or $a(b)$, $a(b')$ the nodes they pass through at height n_0 must be within $2n_0$ of each other. The previous lemma tells us then that for nodes this close together a acts as a fixed translation on the very good branches through these nodes and we have a conflict.

THEOREM 4.1. *Given c and nondegenerate y and y', if $v_n(y, y') < \delta/10$, then*

$$\bar{f}_{[-c\sqrt{n},\, c\sqrt{n}]}(y, y') < \frac{1}{10c^{1.25}} + \frac{C_2}{\sqrt{c}} + \frac{4}{c} < \frac{C_2+5}{\sqrt{c}}.$$

Proof. For this argument weight the nodes at height n_0 uniformly and within each node weight the branches passing through it uniformly. With this weighting the maximum ratio of the weights given two branches passing through nodes in $[-c\sqrt{n}, ..., c\sqrt{n}]$ is $e^{-c^2/2}$. Using this weighting of branches, the weights of nodes at the base and in $[-n+n_0, ..., n-n_0]$ are all equal. Relative to this weighting a fraction $1/10c^{1.25}$ of the nodes are in K_1^c or K'^c_1 and a fraction $C_2/\sqrt{c}$ of the branches through each node of K_1 or K_1' are not very good. A fraction of at most $4e^{-c^2/2}/c$ of the branches through $[-c\sqrt{n}, ..., c\sqrt{n}]$ at height n_0 fail to land in this interval at the base or have an image under a that fails to do so. These branches have a total weight of at most a fraction $4/c$ of the total weight. Hence a fraction in weight of at least

$$1 - \left(\frac{1}{10c^{1.25}} + \frac{C_2}{\sqrt{c}} + \frac{4}{c}\right) > 1 - \frac{C_2+5}{\sqrt{c}}$$

of the branches are nice and land in $[-c\sqrt{n}, ..., c\sqrt{n}]$ and hence at least this fraction of the nodes at the base of the tree are in K_2. ∎

Theorem 2.2 is a restatement of this result. We make a critical observation for our work: For this theorem we only needed

$$\frac{\delta}{10} \leqslant \frac{C_1 C_2 e^{-c^2/2}}{1600 \times 40 \times c^{.75}}$$

but the fraction of branches landing outside of $[-c\sqrt{n}, ..., c\sqrt{n}]$ is less than $2e^{-c^2/2}/c$, a value which grows more slowly in c.

5. INVARIANCE OF ENTROPY

In this section we prove our main result, that if the $[T, T^{-1}]$ endomorphism and the $[S, S^{-1}]$ endomorphism produce isomorphic decreasing sequences of σ-algebras then the entropy of T is equal to the entropy of S. The proof is by contradiction. We assume that $h(S)<h(T)$. Thus in particular S^2 is of finite entropy and has a finite generating partition. Hence we can take S to be the shift map on some $Z=\{1, 2, ..., q\}^{\mathbf{Z}}$ and the points $s \in Z$ are of the form $\{s_i\}_{i \in \mathbf{Z}}$, $s_i \in \{1, ..., q\}$, for all i and just the even terms of the sequence almost surely determines the full sequence, just as we did for T.

We now explain how to construct a *finite code approximation* to the isomorphism Φ. Fix a $c>0$ and let I_n consist of the even values in $[-c\sqrt{n}, c\sqrt{n}]$. Assuming n to be an even perfect square, consider the increasing sequence of finite algebras $\mathscr{C}_n$ generated by sets of the form

$$\{(x, z) \mid (x_0, ..., x_n, z_{-c\sqrt{n}}, ..., z_{c\sqrt{n}}) \in \{-1, 1\}^{[0, n]} \times \{1, ..., q\}^{I_n}\}$$

partitioning $X \times Z$. As these partitions refine to points, the associated algebras will increase to the full algebra of measurable sets.

For a natural number n' and for each $(w, s) \in X \times Z$ consider the n' tree over (w, s). The map Φ takes this tree to the n' tree over $\Phi(w, s)$ by some tree automorphism $a_{n'}(w, s)$. As there are only finitely many automorphisms of a binary tree of height n', this gives a finite partition of $X \times Z$. Namely, partition $X \times Z$ into sets based on which automorphism of the n' tree Φ uses. Call this partition $R=R(n', \Phi)$. For any choice $\varepsilon>0$ there will be an $n \geqslant n'$ so that R can be approximated to within ε by a $\mathscr{C}_n$-measurable partition. That is to say, we can choose an automorphism $a^n_{n'}(w, s)$ of the n' tree over (w, s) depending solely on the value of the first n coordinates of w and the values s_i, $i \in I_n$, and $a^n_{n'}(w, s)$ will agree with $a_{n'}(w, s)$ on all but ε in measure of $X \times Z$.

Now define a new isomorphism Φ_n as follows. Setting $\Phi(w, s)=(u, t)$ we have $\{u_1, ..., u_{n'}\}=a_{n'}(w, s)(\{w_1, ..., w_{n'}\})$. Define the new image point $(u', t')=\Phi_n(w, s)$ by setting $\{u'_1, ..., u'_{n'}\}=a^n_{n'}(w, s)(\{w_1, ..., w_{n'}\})$ and for $i>n'$, $u'_i=u_i$. That is to say, just replace the n' initial terms of u by those terms obtained from the action of $a^n_{n'}(w, s)$. To define t' let $j=\overline{a^n_{n'}(w, s)(\{w_1, ..., w_{n'}\})}-\overline{a_{n'}(w, s)(\{w_1, ..., w_{n'}\})}$ This is the difference between where the original image and the new image branches land. Set $t'=T^j(t)$. It is easy to see that Φ_n is again an isomorphism between the two sequences of σ-algebras. We refer to such a Φ_n as a *finite code approximation* to Φ. Note that for all but ε of $X \times Z$, Φ and Φ_n agree.

As the partitions $\mathscr{C}_n$ increase to the whole σ-algebra we can also construct a $\mathscr{C}_n$-measurable approximation to the partition $\Phi^{-1}(P)$. We call this approximation $\tilde{P}$. This means that we can assign a name to a point (w, s), depending only on the initial n terms of w and the values s_i, $i \in I_n$, and have this name agree with the P-name of $\Phi(w, s)$ on all but ε of $X \times Z$. Note that

Φ and Φ_n may not take $\tilde{P}$ to a partition of $X \times Z$ that depends only on the Z coordinate. If Φ and Φ_n agree at (w, s) and the P-name of $\Phi(w, s)$ agrees with its finite approximation from $\mathscr{C}_n$ then we say (w, s) *codes* n', ε *well* relative to these approximations.

To sketch the path to the conclusion, consider any two points (w, s) and (w', s') for the $[S, S^{-1}]$ endomorphism such that $s_i = s'_i$ for all i in I_n. Suppose $\Phi(w, s) = (u, t)$ and $\Phi(w', s') = (u', t')$. Using a finite code approximation of Φ we will show that there is a natural automorphism $a \in \mathscr{A}_n$ which makes $v_n(t, t')$ small. Theorem 4.1 now implies that $\bar{f}_{[-c\sqrt{n}, c\sqrt{n}]}(t, t')$ is small. This shows that the exponential number of names in the S process must be at least as big as the exponential number of names in the T process, which is a contradiction.

To make this sketch precise, let T and S be 1–1 measure-preserving transformations as described above. For $w \in \{-1, 1\}^n$ define $\bar{w} = \sum_{i=1}^n w_i$. Let Φ be an isomorphism of the decreasing sequences of σ-algebras generated by $[T, T^{-1}]$ and $[S, S^{-1}]$. For any given (w, s), c, and n consider the set

$$E = E_{(w,s), c, n} = \{(w', s') \mid s_{i-\bar{w}} = s'_{i-\bar{w}'} \text{ for all even } i \in [-c\sqrt{n}, c\sqrt{n}]\}.$$

By the Shannon–McMillan theorem for large n we can cover all but ε of the points (w, s) with $2^{(h(S)+\varepsilon)(2c\sqrt{n}+1)}$ of these sets, as we assume we have chosen a generator for S^2.

For any set A define $\Phi^*(A)$ to be the projection of $\Phi(A)$ onto its second coordinate and let F be the set of nondegenerate words for T. We will show that for most of the sets E, $\Phi^*(E) \cap F$ is contained in a small $\bar{f}$ neighborhood. We do this by showing that $\Phi^*(E)$ is contained in a small v_n neighborhood and applying then Theorem 4.1.

LEMMA 5.1. *For any $\varepsilon > 0$, there exists a good set G, with $\mu(G) > 1 - \varepsilon$, c_0, and N_0 such that for any $(w, s) \in G$, $c > c_0$, and $n > N_0$, and setting $\delta = 1/1600c^{1.75}F(c)$, the following property holds: If F is the set of (δ, δ_0) nondegenerate words for T and E is the set associated to (w, s), c, and n and if $(u, t), (u', t') \in \Phi(G \cap E) \cap F$, then*

$$\bar{f}_{[-c\sqrt{n}, c\sqrt{n}]}(t, t') < \varepsilon.$$

Proof. Suppose $\Phi(w, s) = (u, t)$ and $\Phi(w', s') = (u', t')$. There exists a natural automorphism a between the tree over $T^{\bar{u}}t$ and the tree over $T^{\bar{u}'}t'$. This is $a = a_1(Id)(a_2)^{-1}$, where a_1 is the restriction of Φ to the tree of the 2^n preimages of $[T, T^{-1}]^n (w, s)$ and a_2 is the restriction of Φ to the tree of the 2^n preimages of $[T, T^{-1}]^n (w', s')$. Id is the map from the tree of the 2^n preimages of $[T, T^{-1}]^n (w', s')$ to the tree of the 2^n preimages of $[T, T^{-1}]^n (w, s)$ that acts as the identity automorphism on the tree.

Choose c_0 so that $(C_2+5)/\sqrt{c_0}<\varepsilon$. This now defines δ and hence n_0. Let $n'=n_0$ in the previous section. Choose a finite approximation Φ_N to Φ so that the set where the Φ_N codes n', $\delta/4$ well has measure at least $1-\varepsilon$ and set G to be the set where Φ_N codes n', $\delta/4$ well. We will require n to be even larger later. Suppose $(\omega, s)\in G$ and E is the set associated to (ω, s), c, and n. If $(u, t), (u', t')\in \Phi(G\cap E)\cap F$ and Φ_N codes in agreement with Φ at both the branches b and $a(b)$ and both land inside $[-c\sqrt{n}+N, ..., c\sqrt{n}-N]$, then both branches have the same label and the same sequence of 0's and 1's in their last n_0 positions. Assuming that $n>N^2$ such branches include all but a fraction $\delta/2+e^{-(c-1)^2/2}/(c-1)$ of the branches. We can modify a on the remaining branches to assume w.l.o.g. that $f(a)\geqslant n_0$ and hence that

$$v_n(T^{\bar{u}}t, T'^{\bar{u}}t')\leqslant \delta.$$

We conclude then that if n is sufficiently large

$$\bar{f}_{[-c\sqrt{n}, c\sqrt{n}]}(T^{\bar{u}}t, T'^{\bar{u}}t')<\frac{C_2+5}{\sqrt{c}}\leqslant \varepsilon. \quad \blacksquare$$

THEOREM 5.1. *If the decreasing sequence of σ-algebras generated by the* $[T, T^{-1}]$ *endomorphism is isomorphic to the decreasing sequence of σ-algebras generated by the* $[S, S^{-1}]$ *endomorphism then* $h(T)=h(S)$.

Proof. Suppose there was an isomorphism Φ and $h(S)<h(T)$. Consider the generator P described in this section. Excluding a set of P names of small measure, any ε neighborhood in $\bar{f}$ can have a measure of at most

$$\binom{c\sqrt{n}}{\varepsilon c\sqrt{n}}^2 |P|^{\varepsilon c\sqrt{n}}\, 2^{-(1-2\varepsilon)\, h(T)\, c\sqrt{n}}.$$

Thus any set $\Phi^*(G\cap E)\cap F$ can have a measure of at most

$$4n^2\binom{c\sqrt{n}}{\varepsilon c\sqrt{n}}^2 |P|^{\varepsilon c\sqrt{n}}\, 2^{-(1-2\varepsilon)\, h(T)\, c\sqrt{n}}.$$

By the comment above there are at most $2^{(h(S)+\varepsilon)\, c\sqrt{n}}$ of these neighborhoods and they cover all but 2ε of the space. Because ε and n are arbitrary and $h(S)<h(T)$ this is a contradiction. $\blacksquare$

COROLLARY 5.1. *If the* $[T, T^{-1}]$ *endomorphism is isomorphic to the* $[S, S^{-1}]$ *endomorphism then* $h(T)=h(S)$.

Proof. If two endomorphisms are isomorphic then they generate isomorphic decreasing sequences of σ-algebras. Thus Theorem 5.1 implies the corollary. $\blacksquare$

6. REGULARITY CONDITION

There remains a gap in the proof of Theorem 5.1, that is, to define what it means for a branch to be n_0 regular and to develop its properties. To begin we say that a branch b *behaves well* at height h if $|\sum_1^h b_i| < 101\sqrt{h}$. The definition of n_0 regular will imply that if b and b' are n_0 regular branches then, for any $h \geq n_0$, there exists an $h' \in (h^{1/4}, h^{1/2})$ such that both b and b' behave well at height h'.

We will define a sequence $\{h_j\}$ and show not only that there exists such an h', but that it can be chosen to be an element of the sequence h_j. The elements of the sequence h_j will be chosen far enough apart so that we will be able to use the law of the iterated logarithms to prove that the set of branches that behave well at h_j and the set of branches that behave well at h_{j+1} are almost independent. We will then apply the exponential rate of convergence for the weak law of large numbers to show that a large proportion of the branches behave well at more than half of the $h_j \in (h^{1/4}, h^{1/2})$, for all h sufficiently large. Thus, for any two of those branches, we will be able to find at least one height in the interval $(h^{1/4}, h^{1/2})$ where both branches behave well.

Define $h_1 = 100$ and $h_{j+1} = \lfloor h_j(1 + 2\log(\log(h_j)))^2 \rfloor$. Now we show that the set of branches that behave well at h_j and the set of branches that behave well at h_{j+1} are almost independent.

Lemma 6.1. *For any $b_1, ..., b_{h_j}$ such that $|\sum_1^{h_j} b_i| < 2\sqrt{h_j}(\log(\log h_j))$,*

$$P\left(\left|\sum_1^{h_{j+1}} b_i\right| < 101\sqrt{h_{j+1}} \;\middle|\; b_1, ..., b_{h_j}\right) > .9999.$$

Proof. If both $|\sum_{h_j+1}^{h_{j+1}} b_i| < 100\sqrt{h_{j+1}}$ and $|\sum_1^{h_j} b_i| < 2\sqrt{h_j}(\log(\log h_j)) < \sqrt{h_{j+1}}$ are true, then $|\sum_1^{h_{j+1}} b_i| < 101\sqrt{h_{j+1}}$. Since the first event is independent of $b_1, ..., b_{h_j}$, the conditional probability is at least the probability of the first event. By Chebychev's inequality this is at least .9999. ∎

Lemma 6.2. *Between 2^{2^I} and $2^{2^{I+1}}$ there exist at least $2^I/4I$ elements of the subsequence h_j.*

Proof. The ratio h_{j+1}/h_j for any h_j in the interval $[2^{2^I}, 2^{2^{I+1}}]$ is at most $(1 + 2\log(\log(2^{2^{I+1}})))^2$. Thus there are at least k such h_j, where k is the greatest integer such that the inequality

$$2^{2^I}(1 + 2\log(\log(2^{2^{I+1}})))^{2k} < 2^{2^{I+1}}$$

is still true. Now the following calculation proves the lemma,

$$\begin{aligned} 2^{2^I}(1+2\log(\log(2^{2^{I+1}})))^{2k+2} &> 2^{2^{I+1}} \\ (1+2(I+1))^{2k+2} &> 2^{2^I} \\ (2k+2)\log(1+2(I+1)) &> 2^I \\ k &> 2^I/4I. \quad \blacksquare \end{aligned}$$

Now we define

$$G_{h,p} = \left\{ b \;\middle|\; \frac{|\{h_j \mid h^{1/4} < h_j < h^{1/2} \text{ and } |\sum_1^{h_j} b_i| < 101\sqrt{h_j}\}|}{|\{h_j \mid h^{1/4} < h_j < h^{1/2}\}|} > p \right\}.$$

This is the set of branches that behave well on a fraction at least p of the h_j between $h^{1/4}$ and $h^{1/2}$. Now we show that the fraction of branches that follow the law of the iterated logarithms, but are not in $G_{2^{2I},.9}$, is decreasing faster than exponentially in I.

Lemma 6.3. *There exists $C<1$ such that for all I and all $I' \geqslant I$,*

$$\mu\left(G_{2^{2I'+2},.9} \;\middle|\; \left|\sum_1^{h_j} b_i\right| < 2\sqrt{h_j}(\log(\log h_j)) \;\forall h_j > 2^{2^I} \right) < 2\, C^{2^{I'}/4I'}.$$

Proof. For every branch b there exists a subset S of the elements of h_j between $2^{2^{I'}}$ and $2^{2^{I'+1}}$ such that b does not behave well on every $h_j \in S$ and b does behave well on every $h_j \notin S$. If $b \notin G_{2^{2I'+2},.9}$ then the cardinality of S is at least $m/10$, where m is the number of h_j between $2^{2^{I'}}$ and $2^{2^{I'}}$. If S has k elements then Lemma 6.1 implies that the fraction of branches that don't behave well on S is less than $(.0001)^k$. Since the number of subsets with k elements is $\binom{m}{k} < 2^m$, the probability that $b \notin G_{2^{2I'+2}}$, .9 is bounded by

$$\sum_{k=m/10}^{m} 2^m(.0001)^k < 2^m(.0001)^{m/10} \sum_0^{.9m} (.0001)^k < 2(2(.0001)^{.1})^m < 2C^m.$$

The previous lemma says that $m \geqslant 2^{I'}/4I'$, so the lemma is true. $\blacksquare$

Definition 6.1. A branch b is n_0 *regular* if

1. $|\sum_1^{h_j} b_i| < 2\sqrt{h_j}(\log(\log h_j))$ for all $h_j \geqslant n_0$ and
2. $b \in G_{h,.5}$ for all $h \geqslant n_0$.

LEMMA 6.4. *There exists an* N_0 *such that for any* $n_0 > N_0$ *all but* $4/\log n_0$ *of the branches are* n_0 *regular.*

Proof. First assume $n_0 = 2^{2^{I+2}}$ for some I. The rate of convergence to the law of iterated logarithms shows that the fraction of branches that don't satisfy

$$\left|\sum_{1}^{h_j} b_i\right| < 2\sqrt{h_j}(\log(\log h_j)) \qquad \text{for all} \quad h_j \geqslant 2^{2^I} \tag{1}$$

is less than $1/6(2^I)$ for I sufficiently large. Lemma 6.3 implies that, conditioning on (1) holding, the fraction of the branches that do not satisfy

$$b \in G_{2^{2^{I'}+2}, .9} \qquad \text{for all} \quad I' \geqslant I \tag{2}$$

is less than $1/6(2^I)$ for I sufficiently large. Thus (1) and (2) are satisfied for all but $1/3(2^I) < 2/\log n_0$ of the branches.

If b satisfies (1) then b satisfies the first condition in the definition of n_0 regular. Now we show that if b satisfies (2) then b satisfies the second condition in the definition of n_0 regular. If $2^{2^{I'+1}} > h^{1/4} \geqslant 2^{2^{I'}}$ there are at most three times as many h_j between $2^{2^{I'}}$ and $2^{2^{I'+2}}$ as there are between $h^{1/4}$ and $h^{1/2}$. If $b \in G_{2^{2^{I'}+2}, .9}$ and $b \in G_{2^{2^{I'}+3}, .9}$, then b behaves well on at least 90% of the h_j between $2^{2^{I'}}$ and $2^{2^{I'+2}}$. Thus b behaves well on at least 70% of the h_j between $h^{1/4}$ and $h^{1/2}$. So if a branch $b \in G_{2^{2^{I'}+2}, .9}$ for all $I' \geqslant I$ then $b \in G_{h, .7}$ for all h such that $h^{1/4} \geqslant 2^{2^I}$ and $h \geqslant (2^{2^I})^4 = 2^{2^{I+2}} = n_0$. Thus the fraction of branches that are not n_0 regular is less than $2/\log n_0$. Note that if b is n_0 regular then it is n regular for all $n \geqslant n_0$. Thus for arbitrary n_0 the fraction of branches that are not n_0 regular is less than $4/\log n_0$. ∎

LEMMA 6.5. *If* b *and* b' *are* n_0 *regular branches and* $h \geqslant n_0$, *then there exists an* $h' \in (h^{1/4}, h^{1/2})$ *such that* $|\sum_1^{h'} b_i|, |\sum_1^{h'} b'_i| < 101\sqrt{h'}$.

Proof. This follows from the second condition of the definition of n_0 regular. ∎

REFERENCES

1. R. Burton, A non-Bernoulli skew product which is loosely Bernoulli, *Israel J. Math.* **35**, No. 4 (1980), 339–348.
2. J. Feldman, New K-automorphisms and a problem of Kakutani, *Israel J. Math.* **24**, No. 1 (1976), 16–38.
3. J. Feldman and D. Rudolph, Standardness of the decreasing sequences of σ-fields given by certain dyadic endomorphisms, *Fund. Math.* **157**, Nos. 2–3 (1998), 175–189.

4. J. Feldman and M. Smorodinsky, Decreasing sequences of measurable partitions, *Ergodic Theory Dynam. Systems* **20** (2000), 1079–1090.
5. D. Heicklen and C. Hoffman, T, T^{-1} is not standard, *Ergodic Theory Dynam. Systems* **18**, No. 4 (1998), 875–878.
6. C. Hoffman, A zero entropy T such that the (T, Id) endomorphism is not standard, *Proc. Amer. Math. Soc.* **128** (2000), 183–188.
7. S. Kalikow, T, T^{-1} transformation is not loosely Bernoulli, *Ann. of Math.* **115** (1982), 393–409.
8. I. Meilijson, Mixing properties of a class of skew products, *Israel J. Math.* **19** (1974), 266–270.
9. A. Vershik, Decreasing sequences of measurable partitions and their applications, *Dokl. Akad. Nauk SSSR* **193**, No. 4 (1970), 748–751.
10. A. Vershik, On lacunary isomorphism of monotone sequences of partitions, *Funktsional Anal. i Prilozhen.* **2**, No. 3 (1968), 17–21.